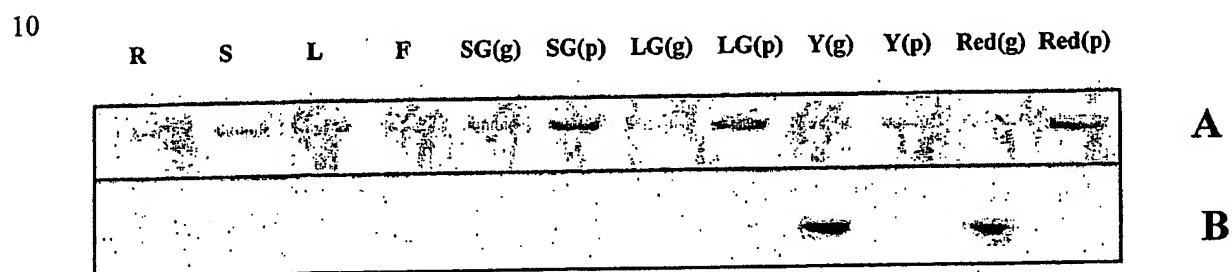
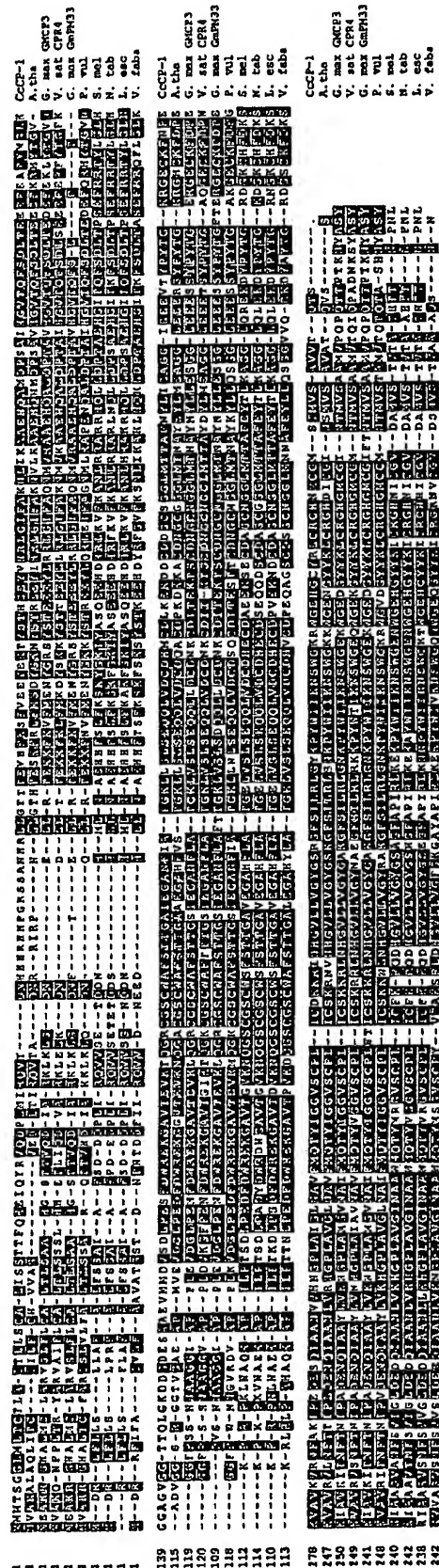


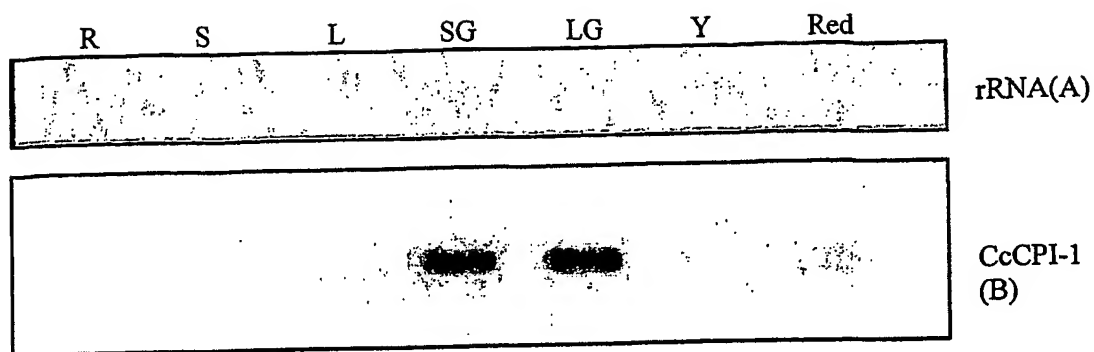
5 **Figure 1:** Northern blot analysis of the expression of the cysteine proteinase (CcCP1) gene in different tissues of *Coffea arabica*.



10  
15 **Figure 2:** Northern blot analysis of the expression of the Cysteine proteinase CcCP-1 gene in different tissues of *Coffea arabica*.

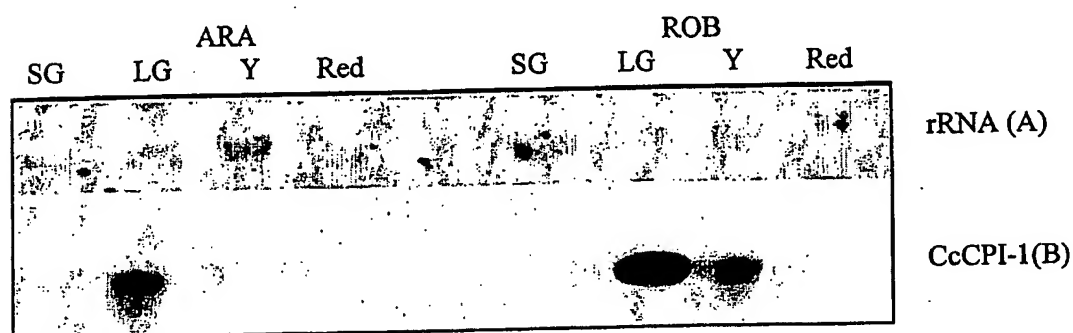


5 **Figure 2A:** Alignment of the full sequence of the protein encoded by CcCP-1 cDNA with other full-length cysteine proteinases available in the NCBI database.



5 **Figure 3:** Northern blot analysis of the expression of the cysteine proteinase inhibitor (CcCPI-1) gene in different tissues of *Coffea arabica*.

10



**Figure 4:** Northern blot analysis of the expression of the cysteine proteinase inhibitor gene (CcCPI-1) at different cherry development stages for *Coffea arabica* (ARA) and *Coffea canephora* (ROB).

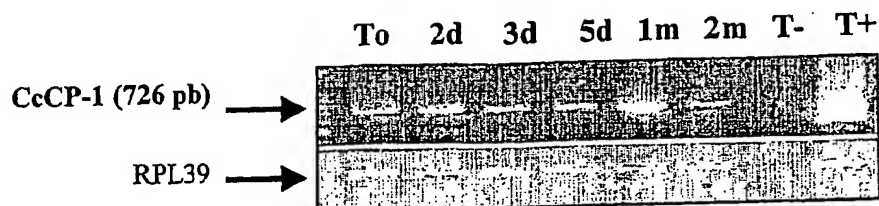


Figure 5. RT-PCR analysis of the expression of *CcCP-1* during *Coffea arabica* grain germination.

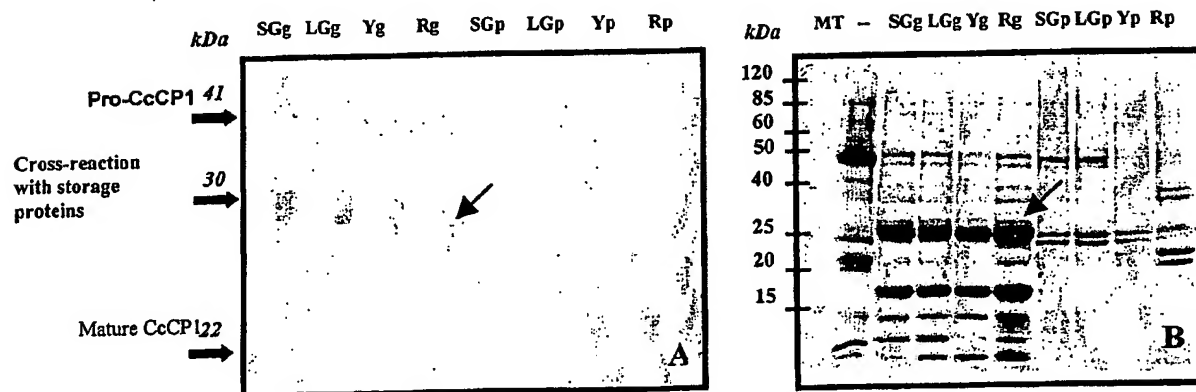


Figure 6: Western-blot analysis of the expression of CcCP1 protein (A).



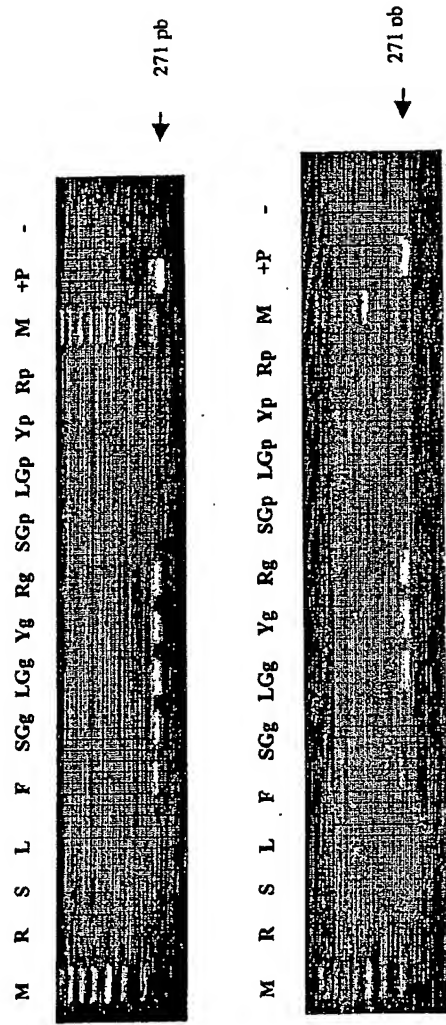


Figure 7: RT-PCR analysis of the expression of CcCPI-1 gene in different tissues of *Coffea arabica* CCA2 (A) and *Coffea robusta* FRT-32 (B).

1	M	A	K	V	G	G	I	S	E	S	K	G	N	-	E	N	S	L	E	I	E	S	L	A	K	F	A	V	D	D	Y	N	K	K	Q	N	A	L	L	E	CcCPI-2
1	M	A	T	I	G	G	I	K	Q	V	E	G	S	-	A	N	S	L	E	V	E	S	L	A	K	F	A	V	E	D	H	N	K	K	Q	N	A	M	L	E	R. obtusifolius
1	M	A	T	V	G	G	I	K	D	S	G	G	S	S	A	N	S	L	E	I	D	E	L	A	K	F	A	V	D	H	Y	N	S	K	E	N	A	L	L	E	D. caryophyllus
1	M	A	T	L	G	G	I	K	E	V	E	E	S	-	A	N	S	V	E	I	D	N	L	A	R	F	A	V	D	D	Y	N	K	K	Q	N	A	L	L	E	M. esculenta
40	F	Q	K	V	I	N	S	K	E	Q	V	V	A	G	T	V	Y	Y	L	T	I	E	V	K	D	G	N	E	K	K	L	Y	E	A	K	V	W	V	K	P	CcCPI-2
40	F	S	K	V	V	N	T	K	E	Q	V	V	A	G	T	M	Y	Y	I	T	L	E	A	T	D	G	G	K	K	K	V	Y	E	A	K	V	W	V	K	P	R. obtusifolius
41	F	Q	R	V	V	N	T	K	E	Q	V	V	A	G	T	I	Y	Y	I	T	L	E	A	T	D	G	G	V	K	K	L	Y	E	A	K	V	W	V	K	P	D. caryophyllus
40	F	K	R	V	V	S	T	K	Q	Q	V	V	A	G	T	M	Y	Y	I	T	L	E	V	A	D	G	G	Q	T	K	V	Y	E	A	K	V	W	E	K	P	M. esculenta
80	W	L	N	F	K	E	V	Q	E	F	K	P	A	A	G	D	T	S	A																					CcCPI-2	
80	W	M	N	F	K	Q	V	Q	E	F	K	L	L	G	D	Q	G	S	T	S																			R. obtusifolius		
81	W	V	N	F	K	E	V	Q	D	F	K	Y	V	G	D	A	S	A																					D. caryophyllus		
80	W	L	N	F	K	E	V	Q	E	F	K	P	I	G	V	A	P	S	D	S	T	A																	M. esculenta		

Figure 8: Optimal alignment of the complete protein encoded by CcCPI-2 cDNA with other homologous full-length cysteine proteinases available in the NCBI.

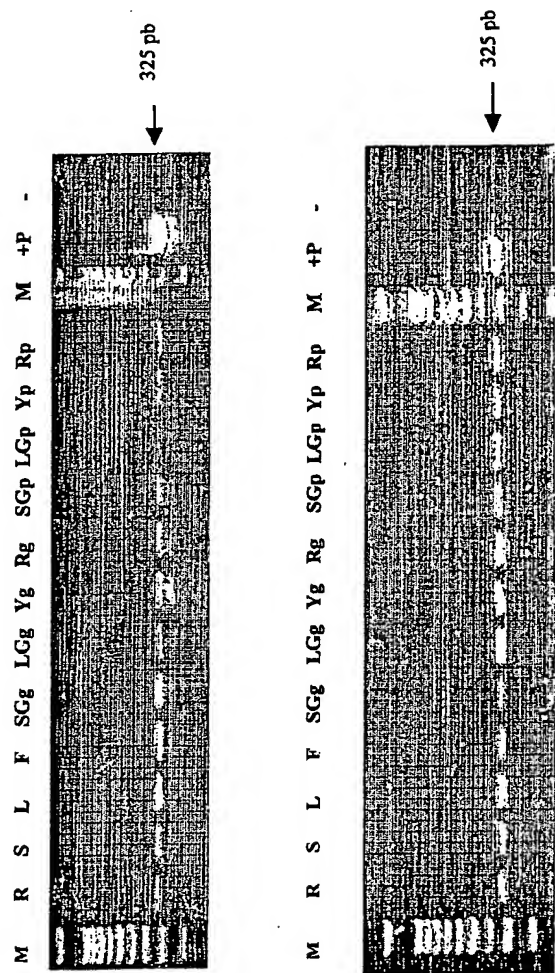


Figure 9: RT-PCR analysis of the expression of CcCPI-2 gene in different tissues of *Coffea arabica* CCA2 (A) and *Coffea robusta* FRT-32 (B).

5

10

15

20

CONFIRMATION COPY



1	M	A	S	A	F	P	H	L	L	L	L	T	T	L	A	A	I	C	L	F	S	D	V	P	S	A	A	L	G	G	R	P	K	D	A	L	V	G	G	W	CcCPL-3	
1	M	N	Q	R	F	C	C	L	I	V	L	-	-	-	-	-	-	-	-	F	L	S	V	V	P	L	L	A	A	G	D	R	-	K	G	A	L	V	G	G	W	Citrus x paradisi
1	M	V	P	K	P	L	S	L	L	L	F	-	-	-	-	-	-	-	-	L	L	L	A	L	S	A	A	V	V	G	G	R	-	K	L	V	A	A	G	G	W	A. deliciosa
1	M	T	S	K	V	V	F	L	L	L	L	-	-	-	-	-	-	-	-	S	L	-	V	V	L	L	L	P	L	Y	A	S	-	A	A	R	V	G	G	W	A. thaliana	
41	S	-	K	A	D	P	K	D	P	E	V	L	E	N	G	K	F	A	I	D	E	H	N	K	E	A	G	T	K	L	E	F	K	T	V	V	E	A	Q	K	CcCPL-3	
33	K	P	I	E	D	P	K	E	K	H	V	M	E	I	G	Q	F	A	V	T	E	Y	N	K	Q	S	K	S	A	L	K	F	E	S	V	E	K	G	E	T	Citrus x paradisi	
33	R	P	I	E	S	L	N	S	A	E	V	Q	D	V	A	Q	F	A	V	S	E	H	N	K	Q	A	N	D	E	L	Q	Y	Q	S	V	V	R	G	Y	T	A. deliciosa	
32	S	P	I	S	N	V	T	D	P	Q	V	V	E	I	G	E	F	A	V	S	E	Y	N	K	R	S	E	S	G	L	K	F	E	T	V	V	S	G	E	T	A. thaliana	
80	Q	V	V	A	G	T	N	Y	K	I	V	I	K	A	L	D	G	T	-	A	S	N	L	Y	E	A	I	V	W	V	K	P	W	L	K	F	K	K	L	T	CcCPL-3	
73	Q	V	V	S	G	T	N	Y	R	L	I	L	V	V	K	D	G	P	-	S	T	K	K	F	E	A	V	V	W	E	K	P	W	E	H	F	K	S	L	T	Citrus x paradisi	
73	Q	V	V	A	G	T	N	Y	R	L	V	I	A	K	D	G	A	-	V	V	G	N	Y	E	A	V	V	W	D	K	P	W	M	H	F	R	N	L	T	A. deliciosa		
72	Q	V	V	S	G	T	N	Y	R	L	K	V	A	A	N	D	G	D	G	V	S	K	N	Y	L	A	I	V	W	D	K	P	W	M	K	F	R	N	L	T	A. thaliana	
119	S	F	R	K	L	P	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	CcCPL-3	
112	S	F	K	P	M	V	K	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Citrus x paradisi	
112	S	F	R	K	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A. deliciosa	
112	S	F	E	P	A	N	N	G	R	F	L	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A. thaliana	

Figure 10: Optimal alignment of the complete protein encoded by CcCPL-3 cDNA with other homologous full-length cysteine proteinases available in the NCBI.

[illegible]

**Figure 11: Optimal alignment of the complete protein encoded by CcCPI-4 cDNA with other homologous full-length cysteine proteinases available in the NCBI.**

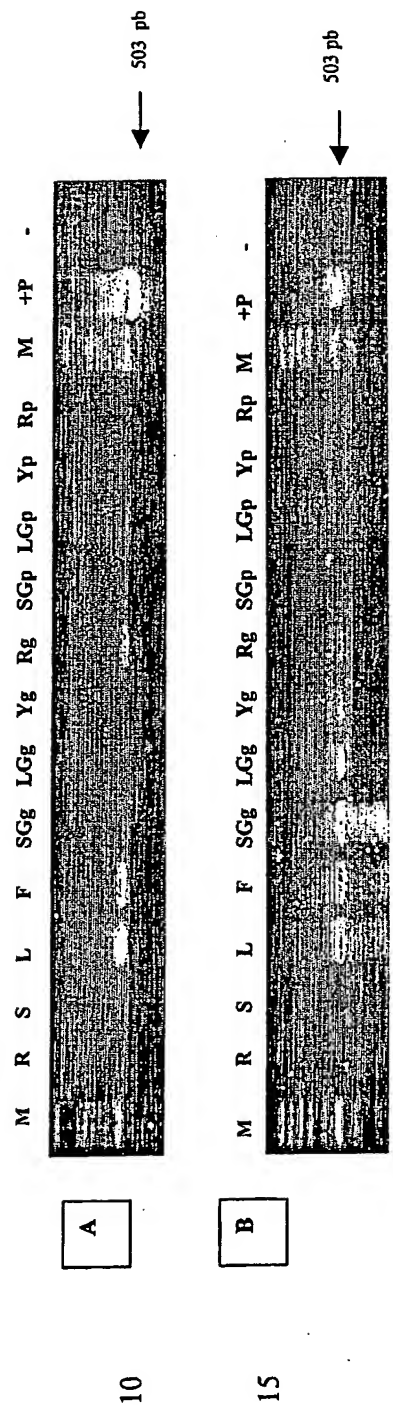


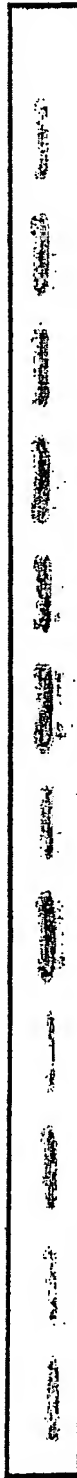
Figure 12: RT-PCR analysis of the expression of CcCPL-4 gene in different tissues of *Coffea arabica* CCA2 (Panel A) and of *Coffea robusta* FRT-32 (Panel B)

12/21

ARABICA

R      S      L      F      SG (G)      SG (P)      LG (G)      LG (P)      Y (G)      Y (P)      Red (G)      Red (P)

rRNA (A)



CcAP-2 (B)

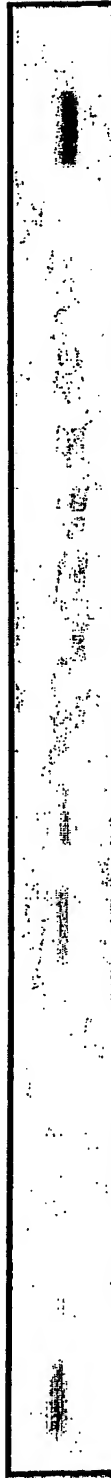


Figure 13: Northern blot analysis of the expression of the aspartic proteinase 2 (CcAP2) gene in different tissues of *Coffea arabica*.

```

1 gcttacatcttaaatcctgatttttatagattcgcccttcgtgaagttcaatcttcgcagtcgctcactaacatttgg
81 agacatacttcgatt ATG AAA ATG GGG AAG GCT TTC CTT TTT GCC GTT GTA TTG GCT GTG ATC
1 M K M G K A F L F A V V L A V I

144 TTA GTG GCG GCT ATG AGC ATG GAG ATC ACA GAA AGA GAT TTG GCT TCT GAG GAA AGC TTG
17 L V A A M S M E I T E R D L A S E E S L

204 TGG GAC TTG TAC GAA AGA TGG AGG AGC CAT CAT ACT GTT TCT CGA GAC CTT TCT GAG AAA
37 W D L Y E R W R S H H T V S R D L S E K

264 CGA AAG CGC TTT AAT GTT TTC AAG GCA AAT GTC CAT CAC ATT CAC AAG GTG AAC CAG AAG
57 R K R F N V F K A N V E H I H K V N Q K

324 GAC AAG CCT TAC AAG CTG AAA CTC AAC AGT TTC GCT GAT ATG ACC AAC CAC GAG TTC AGG
77 D K P Y K L K L N S F A D M T N H E F R

384 GAA TTC TAC AGT TCT AAG GTG AAA CAT TAC CGG ATG CTC CAC GGC AGT CGT GCT AAT ACT
97 E F Y S S K V K H Y R M L H G S R A N T

444 GGA TTT ATG CAT GGG AAG ACT GAA AGT TTG CCA GCC TCC GTT GAT TGG AGA AAG CAA GGA
117 G F M H G K T E S L P A S V D W R K Q G

504 GCC GTG ACT GGC GTC AAG AAT CAA GGC AAA TGT GGT AGC TGT TGG GCA TTT TCA ACT GTG
137 A V T G V K N Q G K C G S C W A F S T V

564 GTT GGA GTC GAG GGA ATC AAC AAA ATC AAA ACA GGC CAA TTA GTT TCT CTG TCC GAG CAA
157 V G V E G I N K I K T G Q L V S L S E Q

624 GAA CTT GTT GAC TGT GAA ACG GAC AAT GAA GGA TGC AAC GGA GGA CTC ATG GAA AAT GCA
177 E L V D C E T D N E G C N G G L M E N A

684 TAC GAG TTT ATT AAG AAA AGT GGG GGA ATA ACA ACT GAG AGG CTA TAT CCC TAC AAG GCA
197 Y E F I K K S G G I T T E R L Y P Y K A

744 AGA GAT GGC AGC TGT GAT TCG TCA AAG ATG AAT GCC CCT GCT GTG ACT ATT GAT GGG CAT
217 R D G S C D S S K M N A P A V T I D G H

804 GAA ATG GTA CCC GCA AAC GAT GAG AAT GCC TTG ATG AAA GCT GTT GCT AAC CAG CCT GTA
237 E M V P A N D E N A L M K A V A N Q P V

864 TCA GTA GCT ATA GAT GCG TCT GGC TCT GAC ATG CAA TTT TAT TCA GAG GGT GTA TAC GCT
257 S V A I D A S G S D M Q F Y S E G V Y A

924 GGA GAC TCG TGT GGC AAT GAG CTT GAT CAT GGC GTG GCG GTC GTC GGC TAC GGG ACT GCT
277 G D S C G N E L D H G V A V V G Y G T A

984 CTT GAC GGT ACT AAA TAC TGG ATA GTG AAG AAC TCA TGG GGA ACA GGA TGG GGA GAA CAG
297 L D G T K Y W I V K N S W G T G W G E Q

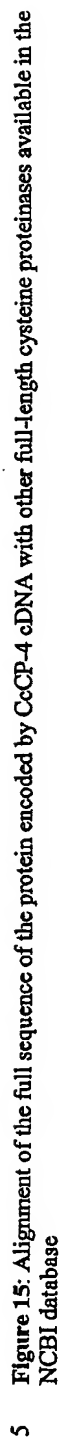
1044 GGC TAT ATC AGG ATG CAA CGT GGT GTT GAT GCT GCT GAA GGC GGA GTT TGT GGG ATA GCA
317 G Y I R M Q R G V D A A E G G V C G I A

1104 ATG GAG GCC TCC TAT CCA CTT AAA TTG TCC TCC CAC AAT CCA AAA CCA TCC CCA CCT AAG
337 M E A S Y P L K L S S H N P K P S P P K

1164 GAC GAC CTC TAG attgatcctcttatatatatacatatatatatatttcagtagattcattgaatttttagttac
357 D D L *
1240 agactacgcgttcTGaagacttagatcatctctagggcatagatttatgtaatcctgctcctgtgatgggttgaataaac
1320 aataagtagtactaataaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

```

**Figure 14:** cDNA sequence and its deduced amino acid sequence of CcCP-4. Lowercase: 5' and 3' non-translated regions; Uppercase: Open reading frame; Bold character: amino acid sequence; \*: stop codon





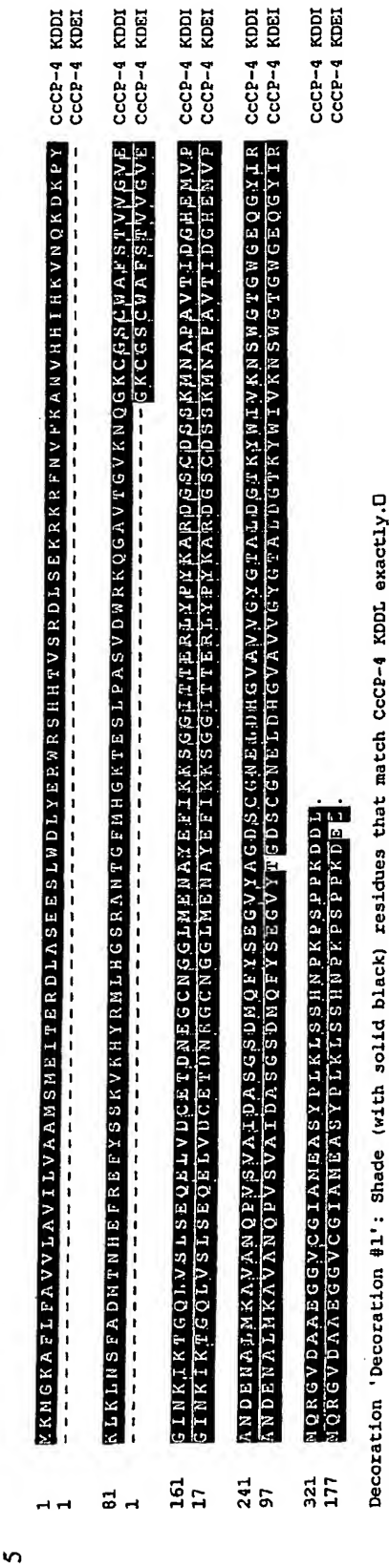
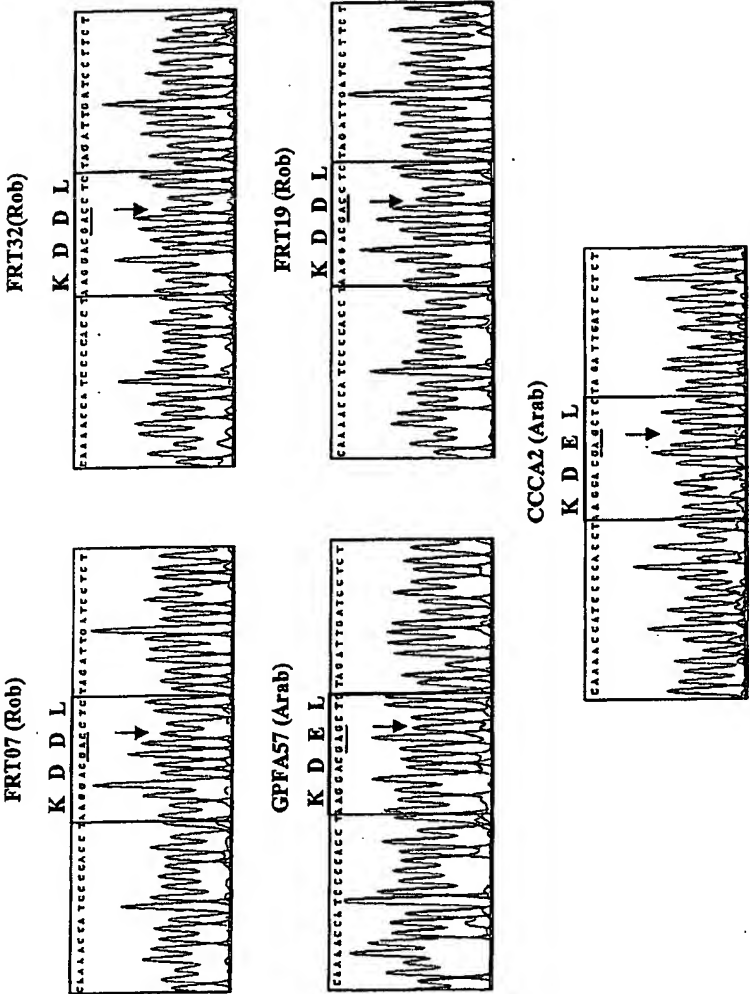


Figure 17. The complete open reading frame of CcCP-4 (KDDL) and the partial open reading frame of CcCP-4 (KDEL).





5 Figure 18. DNA sequence chromatograms for PCR amplified genomic DNA encoding the KDEL/KDDL region of the CcCP-4 gene.

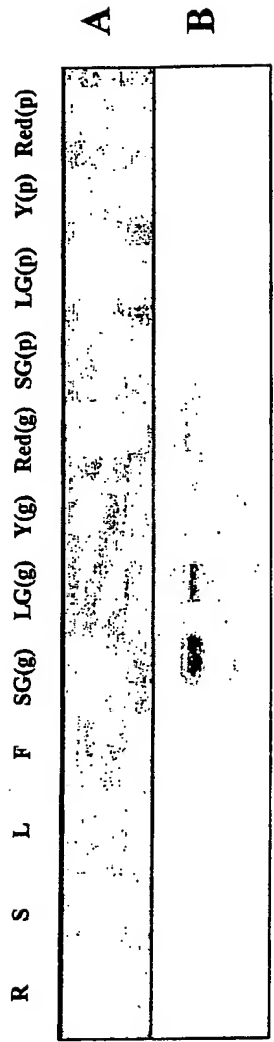


Figure 19. Northern blot analysis of the expression of the Cysteine proteinase CcCP-4 gene in different tissues of *Coffea arabica*.

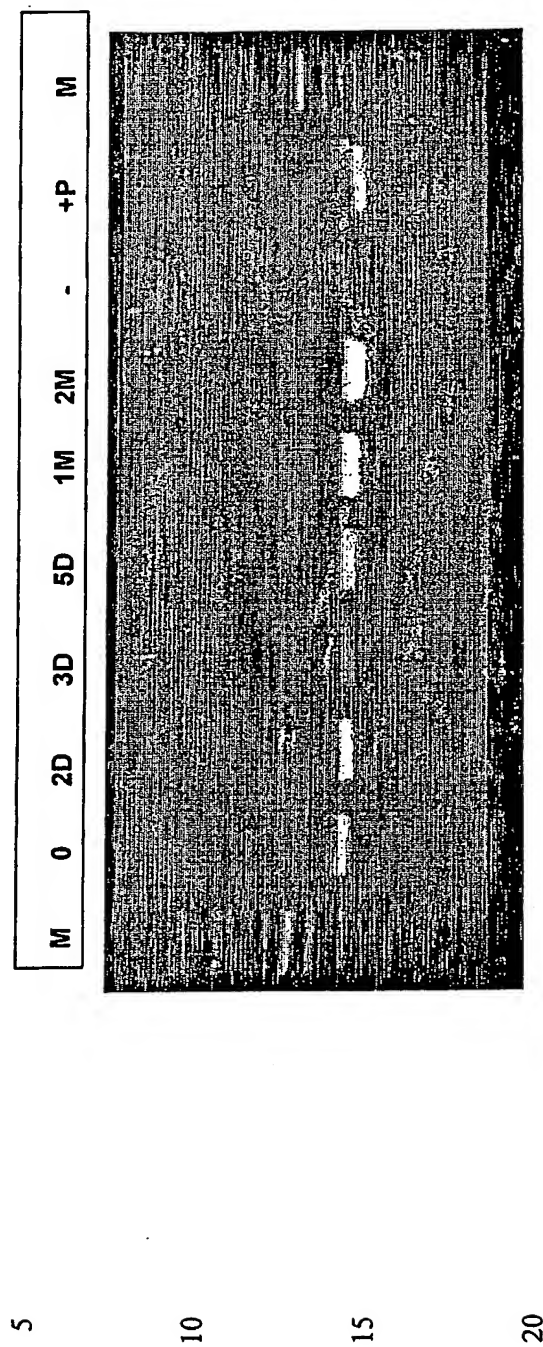


Figure 20. RT-PCR analysis of the expression of CcCP-4 in the whole grain during germination.

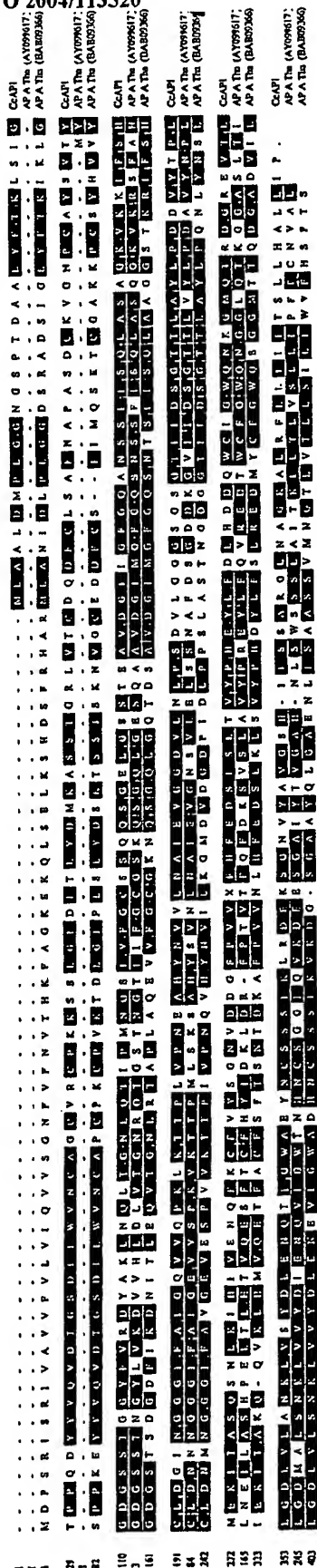


Figure 21: Optimal alignment of the complete protein encoded by CcAP-1 cDNA with other homologous full-length aspartic proteinase sequences available in the NCBI.

1 MERRY LWAALF VLGAIVCSLFLPSEGG -- LKRI SLKKKKPLDIQSTI RAAKLAHLESTHGAGARKEN - DNNL CcAP2  
 1 MGQKHLVITVFCWALTCSLLESFSFG -- ILRI GLKKRPDLDSI NAARKAREGLRSVRPMMGAHQEI G.max  
 1 MGRKHLCAALLLWALVCTALPAAYSNNLLRVLGLKKRPLDLESIKAARKARLGGKYGKGVN -- -- KKKL I. batatas  
 1 MDKKHLCAALLLWALVCTALPAAYSNNLLRVLGLKKRPLDLESIKAARKARLGGKYGKGVN -- -- KKKL I. batatas  
 1 MGRNRLWVIECFICALISCFESTSADG -- LVRI GLKKRPQSDNSIRAVIRAKAGHN - QGILKRKF - QYSF N. alata N  
 66 GSSN - EDILPLKKNYLDDAQYVGEIGLIGTPPOKAFVJEDT GSSNLLWVPSAKCYFSLIACW LHSKYAKKSS CcAP2  
 67 GKSKGEDIPLKKNYLDDAQYVGEIGLIGTPPOKAFVJEDT GSSNLLWVPSAKCYFSLIACW LHSKYAKKSS CcAP2  
 65 GDSKD - EGIVSLNNYLDDAQYVGEIGLIGTPPOKAFVJEDT GSSNLLWVPSAKCYFSLIACW LHSKYAKKSS I. batatas  
 66 GDSKD - IYKVP LKKNYLDDAQYVGEIGLIGTPPOKAFVJEDT GSSNLLWVPSAKCYFSLIACW LHSKYAKKSS I. batatas  
 65 GDSKD - TDIVY LKKNYLDDAQYVGEIGLIGTPPOKAFVJEDT GSSNLLWVPSAKCYFSLIACW LHSKYAKKSS N. alata N  
 133 TYTAKGKSCIRYGSGLISGFSQDNVEVGD LKVKDQVFIKASREGSLTFLAIKEDGILGLGFQEI AV CcAP2  
 135 THVNGTSCIRYKINYGTCISGFSQDNVEVGD LKVKDQVFIKASREGSLTFLAIKEDGILGLGFQEI AV CcAP2  
 132 TYTQKSCIRYKINYGTCISGFSQDNVEVGD LKVKDQVFIKASREGSLTFLAIKEDGILGLGFQEI AV I. batatas  
 133 TYTQKSCIRYKINYGTCISGFSQDNVEVGD LKVKDQVFIKASREGSLTFLAIKEDGILGLGFQEI AV I. batatas  
 132 TYTQKSCIRYKINYGTCISGFSQDNVEVGD LKVKDQVFIKASREGSLTFLAIKEDGILGLGFQEI AV N. alata N  
 201 DNMVPVWYRMV DQGLVDEQVFSFWLNRDPNAE DGGELVFGGVGI NHEIKGKHTIYVPTONGYMOFKMGD CcAP2  
 203 ENAVPVVMEKMEVQKLISENVEFEMNLNGDPNAE DGGELVFGGVGI NHEIKGKHTIYVPTONGYMOFKMGD G.max  
 200 ENAVPVVMEKMEVQKLISENVEFEMNLNGDPNAE DGGELVFGGVGI NHEIKGKHTIYVPTONGYMOFKMGD I. batatas  
 201 GHTFVWYRMV DQGLVDEQVFSFWLNRDPNAE DGGELVFGGVGI NHEIKGKHTIYVPTONGYMOFKMGD I. batatas  
 200 GHTFVWYRMV DQGLVDEQVFSFWLNRDPNAE DGGELVFGGVGI NHEIKGKHTIYVPTONGYMOFKMGD N. alata N  
 269 FLIGRVSTGFCGCGCAALVDSGTSLLAGFTT VVTOINHAIGNEGVSVECKEIVSOYGE LHWDLVLSG CcAP2  
 271 FEVGGVSTGFCGCGCAALVDSGTSLLAGFTT VVTOINHAIGNEGVSVECKEIVSOYGE LHWDLVLSG G.max  
 268 FLIGRVSTGFCGCGCAALVDSGTSLLAGFTT VVTOINHAIGNEGVSVECKEIVSOYGE LHWDLVLSG I. batatas  
 269 FLIGRVSTGFCGCGCAALVDSGTSLLAGFTT VVTOINHAIGNEGVSVECKEIVSOYGE LHWDLVLSG I. batatas  
 268 FLIGRVSTGFCGCGCAALVDSGTSLLAGFTT VVTOINHAIGNEGVSVECKEIVSOYGE LHWDLVLSG N. alata N  
 337 VLPRVCKOAGLCLPRGAONEMAY LNSVVDLEKKEASVGSSEMCCTACEMAVVMONQLKOQSTKEKV CcAP2  
 339 VKPRDDICSQVGLCSKAKHQBKSAQTEHVT EKLQ - ELAARQTPLCSGQMLV LWTQNLKQKATKDRV G.max  
 336 LRAQVCS EIGLCFLNGAWHESITKTVEKEA - EG - NLTSNPLCTGEMAVIWLQNLKQKATKDRV I. batatas  
 337 IRPDQVCSQAGLCFLNGAWHESITKTVEKEA - EG - NLTSNPLCTGEMAVIWLQNLKQKATKDRV I. batatas  
 336 VOPKICSQLALCFN - DAQFLSISIKTVIERENRKNSSVALDFLCTACENAVVMONQLKOQSTKEKV N. alata N  
 405 LA VYNQCESIPSPMGES IYDCNLSLT LRVNVSFTJLGRSFEETLKEEVLRTIGEGFAEVCI SIGE HAMMDV CcAP2  
 406 FVYNQCESIPSPMGES IYDCNLSLT LRVNVSFTJLGRSFEETLKEEVLRTIGEGFAEVCI SIGE HAMMDV G.max  
 402 FEYDOLCEKLPSPMGES IYDCNLSLT LRVNVSFTJLGRSFEETLKEEVLRTIGEGFAEVCI SIGE HAMMDV I. batatas  
 404 LEYDOLCEKLPSPMGES IYDCNLSLT LRVNVSFTJLGRSFEETLKEEVLRTIGEGFAEVCI SIGE HAMMDV L. esculen  
 403 LNYINELCDSLPSPMGES IYDCNLSLT LRVNVSFTJLGRSFEETLKEEVLRTIGEGFAEVCI SIGE HAMMDV N. alata N  
 473 PPRRGPIVVLGDVFMGVYHIVFDYGNLRN GFARAAV CcAP2  
 474 PPRKGLVILGDVFMGVYHIVFDYGNLRN GFARAAV G.max  
 470 PPRKGLVILGDVFMGVYHIVFDYGNLRN GFARAAV I. batatas  
 472 PPRKGLVILGDVFMGVYHIVFDYGNLRN GFARAAV L. esculen  
 471 PPRKGLVILGDVFMGVYHIVFDYGNLRN GFARAAV N. alata N

Figure 22: Optimal alignment of the complete protein encoded by CcAP-2 cDNA with other homologous full-length aspartic proteinase sequences available in the NCBI